AD-A057 398 STANFORD UNIV CALIF SYSTEMS OPTIMIZATION LAB F/G 12/1

A LOCALLY QUADRATICALLY CONVERGENT ALGORITHM FOR COMPUTING STAT--ETC(U)

MAY 78 B C EAVES DAAG29-78-G-0026

UNCLASSIFIED SOL-78-13 ARO-15254.2-M NL

| OF |

ADA

057398

END

DATE

FILMED

9 78

DDC

LEVEL II

(4)

A LOCALLY QUADRATICALLY CONVERGENT ALGORITHM FOR COMPUTING STATIONARY POINTS

TECHNICAL REPORT

B. CURTIS EAVES

MAY 1978

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

This research was supported in part by Army Research Office-Durham Contract DAAG 29-78-G-0026, Department of Energy Contract EY-76-S-03-0326 PA #18 and National Science Foundation Grant MCS 77-05623.

Abstract

Stationary points for a system with a convex polyhedral set and a continuously differentiable function with positive definite derivatives are computed by iteratively solving the linearized problem. The procedure is shown to be a mixing of a finite number of Newton methods and to have a local convergence rate which is quadratic. The stationary point problem is of the type arising from the PIES energy model.

A LOCALLY QUADRATICALLY CONVERGENT
ALGORITHM FOR COMPUTING STATIONARY POINTS

1. Introduction

Let Y and Z be two sets in R^n and let $g : Z \to R^n$ be a function. A point y_* in Y and in the domain of g is defined to be a stationary point or solution to the variational problem (Y,g) if

$$(y_* - y)\, g(y_*) \leq 0$$

for all y in Y.

In this paper we are concerned with the computation of stationary points of $(X,(e,f)) = (X,e,f)$ where X in $R^{\ell+m} = R^n$ is of form $\{(u,v) : Au + Bv \leq C\}$ and $(e,f)(u,v) = (e,f(v))$. The vector e is any element of R^{ℓ}. Let V be the set of v for which there is a (u,v) in X. The function $f : V \to R^m$ is assumed to be continuously differentiable (C^1) and to have derivatives that are positive definite (p.d.) but not necessarily symetric.

The algorithm defined below is shown to be a mixing of a finite number of Newton methods. If v_0 is near v_* where (u_*,v_*) is a stationary point of (X,e,f), then the iterates v_k generated by the algorithm are shown to converge at a superlinear rate to v_*. If, in addition, the derivative f' of f is Lipschitz continuous at v_*, the convergence rate is quadratic.

For convenience we shall refer to $\bar{v}$ as a v-stationary point of (X,e,f), if there is a $\bar{u}$ such that $(\bar{u},\bar{v})$ is a stationary point of (X,e,f). Also for each $\bar{v}$ we define the function $f\bar{v} : R^m \to R^m$ by

$$f\bar{v}(v) = f(\bar{v}) + f'(\bar{v})(v-\bar{v})$$

Regard $f\bar{v}(v)$ as an approximation of $f(v)$ from $\bar{v}$.
Define $G : V \to V$ by setting $G(\bar{v})$ to be the v-stationary point of $(X,e,f\bar{v})$. The algorithm is defined by $v_k = G^k(v_0)$ for $k = 1, 2, \ldots$ where v_0 is selected arbitrarily in V. That is to say, the kth iterate is $v_k = G(v_{k-1})$ where v_{k-1} is the $(k-1)$th iterate.

Executing step k, that is, computing $G(v_{k-1})$ is a matter of solving a linear complementary problem. Procedures for solving such a problem can be found in Cottle [2] and Lemke [9], however, of special interest here is [3]. The effort required to evaluate $G(v_k)$ is approximately that of solving or post-optimizing one linear program with the simplex method where the matrix, including the basis, is $n' + n$ by $2(n' + n)$ and (A,B) is n' by n.

The motivation for this paper was due in part from an interest in the PIES energy model and algorithms thereof, see Hogan [5 , 6]. Let us briefly describe this model and how it can be formulated as a stationary point problem. First, on the demand side of the model we have a function $\mathcal{Q}$ of prices p. At the vector of prices p there is a vector of demands $\mathcal{Q}(p)$; it is assumed that $\mathcal{Q}$ is C^1 and its

derivatives are negative definite but not necessarily symetric. On the supply side we have the linear program

$$\begin{aligned} \underset{x}{\text{minimize:}} \quad & cx \\ \text{subject to:} \quad & Ax = a \\ & Bx = q \qquad x \geq 0 \end{aligned}$$

where q is the vector of quantities to be produced. Roughly, an equilibrium of the model is defined to be a price p_* such that the marginal cost of producing at levels $\mathcal{Q}(p_*)$ is p_*. More precisely, if the linear program is solved for $q = \mathcal{Q}(p_*)$, then p_* must be a set of optimal dual prices for q. In other words, we desire a p_* for which there is a pair (λ_*, p_*) which solves the linear program (1).

(1)
$$\begin{aligned} \underset{(\lambda,p)}{\text{maximize:}} \quad & \lambda a + p\,\mathcal{Q}(p_*) \\ \text{subject to:} \quad & \lambda A + pB \leq c \end{aligned}$$

By setting $e = -a$ and $f = -\mathcal{Q}$ it is clear that this is the stationary point problem where f has p.d. derivatives. If the inverse demand function $\mathcal{P} = \mathcal{Q}^{-1}$ is available we can state the problem as that of finding a q_* for which there is a (x_*, q_*) solving the linear program (2).

(2)
$$\begin{aligned} \underset{x,q}{\text{minimize:}} \quad & cx - \mathcal{P}(q_*)\,q \\ \text{subject to:} \quad & Ax = a \\ & Bx - q = 0 \qquad x \geq 0 \end{aligned}$$

Setting $e = c$ and $f = -\mathscr{P}$ we again have a stationary point problem with p.d. derivatives for f . It is easy to show that p_* solves (1) if and only if q_* solves (2) where $q_* = \mathscr{Q}(p_*)$.

The PIES algorithm proceeds in an effort to find q_* by solving at the $(k+1)$th iteration the linear program

$$\begin{aligned} \underset{x,q}{\text{minimize:}} \quad & cx - C(q_k,q) \\ \text{subject to:} \quad & Ax = a \\ & Bx - q = 0 \qquad x \geq 0 \end{aligned}$$

for q_{k+1} where $C(q_k,q)$ as a function of q is a convex piecewise linear function which approximates the integral

$$\sum_{i=1}^{h} \int_0^{q^i} \pi_i(z;q_k)\, dz$$

where

$$\pi_i(z;q_k) = \mathscr{P}(q_k^1, \ldots, q_k^{i-1}, z, q_k^{i+1}, \ldots, q_k^h)$$

and q^i is the ith component of q .

The algorithm proposed here for computing q_* would proceed at the $(k+1)$th iteration by solving for a q^{k+1} such that (x_{k+1}, q_{k+1}) solves the linear program

$$\underset{x,q}{\text{minimize:}} \quad cx - \mathscr{P}_{q_k}(q_{k+1})\, q$$

$$\text{subject to:} \quad Ax = a$$

$$Bx = q \qquad x \geq 0$$

Clearly the subproblems of the PIES algorithm are considerably easier to solve than those of our algorithms; roughly, we are comparing a n' by $n + n'$ linear program with a $(n'+n)$ by $2(n'+n)$ linear program. In both cases one can use the solution of the last iteration to initiate the current iteration.

The PIES algorithm is reported to perform remarkably well [5,6]. However, on a theoretical level it has to date evaded a complete analysis. A number of pertinent investigations include Ahn [1], Irwin [7], Thrasher [22], and Tse [24]. In particular, Ahn has demonstrated that for a linear demand function the convergence is global with a linear rate of convergence.

Looking now in other quarters of the literature we first note the work of Mathiesen [10,11,12]. He observed that problems of type (1) and (2) with linear $\mathscr{Q}$ or $\mathscr{P}$ could be cast into the form of the linear complementarity problem and solved, if there is a solution, by Lemke's method. Solving such a problem is step k of the algorithm of this paper, as already mentioned.

Next we observe that the algorithm of this paper is closely related to the scheme of Wilson [25] for maximizing a nonlinear objective subject to nonlinear constraints; his scheme proceeds by solving a sequence of quadratic programs. Wilson did not have a proof, however,

in recent years there have been a number of arguments validating this type of algorithm wherein the schemes are shown to be like Newton's method in the tail, see, for example, Han [4], Palomares and Mangasarian [14], Powell [15,16], Tapia [20,21] and Robinson [17,18]. Often these schemes utilize only the derivatives of the objective and not the objective itself. Hence, these schemes could be used as well for stationary point problems. Following this reasoning, one obtains after suitable specialization and relaxation the scheme of this paper. Our analysis here does not use multipliers and we have been able to drop the regularity conditions required of these procedures.

The studies [4 ,14,15,16,17,18,20,21] focus on the fact that superlinear convergence is retained, if the quadratic form at each iteration only approximates the Hessian (our derivative). The essential idea is that the computation of the Hessian at each iteration is too expensive and that an approximation serves almost as well. No doubt the algorithm of the paper could be improved in this sense. Perhaps the PIES algorithm is related to the algorithm of the paper in this regard.

Also of interest are the global schemes of Rockafellar [19] and Todd [23] for solving the stationary point problem; however, both schemes are quite different than what is contemplated here. One could invision using the algorithm proposed here as a tail routine for these global procedures.

2. Newton's Method

Let us consider the application of Newton's method to solve the system

$$(3) \qquad H(u,v) = Eu + F(v) = 0$$

of n equations in n variables (u,v) where Eu is linear in u in R^{ℓ} and F is C^1 with domain V.

Given the estimate (u_0,v_0) in $R^{\ell} \times V$ the Newton iterate is any (u_1,v_1) in R^n which solves

$$Eu_0 + F(v_0) + E(u_1-u_0) + F'(v_0)(v_1-v_0) = 0$$

that is, which solves

$$(4) \qquad Eu_1 + Fv_0(v_1) = 0$$

Note that (u_1,v_1) depends upon v_0 but not u_0 and that such a (u_1,v_1) may not exist. Let us assume that (u_*,v_*) solves (3) and that any solution (u,v) to

$$Eu + Fv_*(v) = 0$$

has $v = v_*$. Finally, let us consider only those v_0 for which

v falls in V. We argue that Newton's method in this setting has the following convergence properties. First, there is a $\delta > 0$ and $\beta < 1$ such that if $\|v_0 - v_*\| \leq \delta$ then $\|v_1 - v_*\| \leq \beta\|v_0 - v_*\|$. Second, β can be made arbitrarily small by making δ sufficiently small. Third, if if F' is Lipschitz continuous, then there is a γ so that $\|v_1 - v_*\| \leq \gamma\|v_0 - v_*\|^2$ for sufficiently small δ.

Observe we do not make the customary assumption that the matrix $(E,F'(v_*))$ is nonsingular, however, we have required that $F'(v_*)$ be of full column rank through the uniqueness condition.

Let $d_1, \ldots d_h$ be an orthogonal basis for the null space $\{u : Eu = 0\}$ of E. Let D be the matrix

$$\begin{pmatrix} d_1 \\ \cdot \\ \cdot \\ \cdot \\ d_h \end{pmatrix}$$

Then for any u there exist a unique $\bar{u}$ such that

$$E\bar{u} = Eu$$

$$D\bar{u} = 0$$

For any matrix M let M_α denote the submatrix of rows indexed by α. Select α so that the matrix

$$\begin{pmatrix} E_\alpha & F'_\alpha(v_*) \\ D & 0 \end{pmatrix}$$

is nonsingular. Consider the system (5).

$$(5) \qquad \begin{pmatrix} E_\alpha \\ D \end{pmatrix} u + \begin{pmatrix} F_\alpha(v) \\ -Du_* \end{pmatrix} = 0$$

Of course, (u_*,v_*) is a solution to (5). Now we apply Newton's method to the system (5) at (u_*,v_*). Given v_0 in V the system

$$(6) \qquad \begin{pmatrix} E_\alpha \\ D \end{pmatrix} u + \begin{pmatrix} F_\alpha v_0(v_1) \\ -Du_* \end{pmatrix} = 0$$

is solved for (u_1, v_1). If we assume that v_1 lies in V, the usual argument for the convergence of Newton's method now obtains, namely, a nonsingular derivative at the solution and continuous differentiability, see Ortega [13], for example. Since the computation of (u_1,v_1) does not involve u_0 we may assume that $u_1 = u_*$ and conclude: First, there is a $\delta > 0$ and $\beta < 1$ such that if $\|(u_*,v_0) - (u_*,v_*)\| \leq \delta$ then $\|(u_1,v_1) - (u_*,v_*)\| \leq \beta\|(u_*,v_0) - (u_*,v_*)\|$. Second, β can be made arbitrarily small by making δ sufficiently

small. Third, if F'_α is Lipschitz continuous, then there is a γ so that $\|(u_1,v_1) - (u_*,v_*)\| \leq \gamma\|(u_*,v_0) - (u_*,v_*)\|^2$ for sufficiently small δ .

Recognizing that if v_1 solves (4) then it solves (6), we have established the convergence properties of the process (4).

3. Uniqueness and Existence

In this section we cite a number of uniqueness and existence results which are necessary for or clarify the subsequent development. We shall use the box norm, that is, $\|x\| = \max|x_i|$.

The results of this section are about the structures X and $\bar{X}$ in $R^{\ell+m}$, $V = \{v : (u,v) \in X\}$, $e \in R^{\ell}$, $f : V \to R^m$, and $g : X \to R^{\ell+m}$. If X is of form $\{(u,v) : Au + Bv \leq c\}$, then we shall refer to X as a cell.

We define f to be strictly monotone if

$$(\bar{v} - v)(f(\bar{v}) - f(v))$$

is positive for all distinct $\bar{v}$ and v in V, and we define f to be strongly monotone if for some $\alpha > 0$

$$(\bar{v} - v)(f(\bar{v}) - f(v)) \geq \alpha \|\bar{v} - v\|^2$$

for all v and $\bar{v}$ in V . We say the derivatives of f are p.d. if $y\, f'(v)y$ is positive for all (y,v) in $R^m \times V$ with $y \neq 0$, and we say the derivatives of f are uniformly p.d. if for some $\alpha > 0$

$$y\, f'(v)y \geq \alpha \|y\|^2$$

for all (y,v) in $R^m \times V$.

__Lemma 1__: Assuming X is convex and f is C^1 with p.d. or uniformly p.d. derivatives, then f is strictly monotone or strongly monotone, respectively.

__Proof__: The results follow from the fundamental theorem of integral calculus, that is,

$$f(x) - f(y) = \int_0^1 f'(x + t(y - x))(y - x)\, dt \qquad \boxtimes$$

Also, we note that if f is C^1 and strongly monotone, then the derivatives of f are uniformly positive definite. If the derivatives of f are strictly monotone and V is compact, then the derivatives are strongly monotone. The next result is patterned after Karamardian [8].

__Lemma 2__: If f is strictly monotone, then (X,e,f) has at most one v-stationary point v_* .

__Proof__: Let (u_0,v_0) and (u_*,v_*) be stationary points. Then

$$(u_* - u_0)\, e + (v_* - v_0)\, f(v_*) \leq 0$$

$$(u_0 - u_*)\, e + (v_0 - v_*)\, f(v_0) \leq 0$$

hence

$$(v_* - v_0)\, (f(v_*) - f(v_0)) \leq 0 \ ,$$

and hence $v_0 = v_*$. $\boxtimes$

For v in V define $U(v)$ to be the set of all optimal solutions u of the program

$$\begin{aligned} &\underset{u}{\text{minimize:}} \quad eu \\ &\text{subject to:} \quad (u,v) \in X \end{aligned}$$

<u>Lemma 3</u>: If f is strictly monotone and v_* is the v-stationary point, then the set of stationary points of (X,e,f) is $U(v_*) \times \{v_*\}$. $\boxtimes$

<u>Lemma 4</u>: If f is strictly monotone and some stationary point of $(\bar{X},e,f)$ is also a stationary point of (X,e,f), then all stationary points of $(\bar{X},e,f)$ are stationary points of (X,e,f) .

<u>Proof</u>: If (u_1,v_1) and (u_2,v_1) are both stationary points of $(\bar{X},e,f)$ then $eu_2 = eu_1$ and hence, (u_2,v_1) is a stationary point of (X,e,f) . $\boxtimes$

<u>Lemma 5</u>: If (u,v) is a stationary point of (X,g) and (u,v) is in $\bar{X}$, then (u,v) is a stationary point of $(X \cap \bar{X},g)$. $\boxtimes$

For X closed and convex define $\pi : R^n \to X$ to be the projection to X, namely, $\pi(x)$ is the point in X closest to x; observe that π is continuous.

Lemma 6: If X is closed and convex, then x is a stationary point of (X,g) if and only if $\pi(x - g(x)) = x$. ⊠

Lemma 7: Let X be convex. If x is a stationary point of $(X \cap \bar{X}, g)$ and x is interior to $\bar{X}$, then x is a stationary point of (X,g) .

Proof: x is the point in $X \cap \bar{X}$ closest to $(x - g(x))$. Thus, x is the point in X closest to $(x - g(x))$. ⊠

By a face φ of a convex set X we mean the meet of X with a supporting hyperplane or the set X itself.

Lemma 8: If X is convex, x is a stationary point of (X,g), and x lies in the relative interior of a face φ of X, then x is a stationary point of $(\bar{X},g)$ where $\bar{X}$ is the affine hull of φ.

Proof: If $\bar{x}$ is in X, then $\bar{x} = x + \lambda(y - x)$ where y is in φ, consequently, $(x - \bar{x})\ g(x) = \lambda(x - y)\ g(x) \leq 0$. ⊠

Lemma 9: If X is a cell and (X,e,f) has a stationary point, then

$$Au \leq 0 \qquad eu < 0$$

has no solution.

Proof: It follows immediately from the definitive of a stationary point. ⊠

Clearly $Au \leq 0$ with $eu < 0$ has no solution, if and only if for each sequence (u_k, v_k) in X for which eu_k tends to $-\infty$, then v_k tends to ∞, that is, $\|v_k\|$ tends to $+\infty$. Obviously, the PIES models (1) and (2) possess this property or no solution would exist; in any case, one can give a simple economic argument that such should be the case.

The next lemma guarantees that each step of our algorithm, namely, the evaluation of $G(v_k)$, can be executed. Let $\bar{X} = \{(u,v) : \bar{A}u + \bar{B}v \leq \bar{c}\}$.

Lemma 10: If $\bar{X}$ is a cell, $\bar{A}u \leq 0$ with $eu < 0$ has no solution, and $f'(v)$ is p.d., then (X,e,fv) has a stationary point.

Proof: According to [3] if Lemke's algorithm cannot compute a stationary point of $(\bar{X},e,fv)$ then there is a ray $\{(u_1,v_1) + \theta(u_2,v_2) : \theta \geq 0\}$ in $\bar{X}$ with $u_2\, e + v_2\, fv(v_1 + \theta v_2)$ negative, that is, $u_2\, e + v_2\, f(v) + v_2 f'(v) v_1 + \theta v_2 f'(v) v_2$ negative for all θ. Since $f'(v)$ is p.d., $v_2 = 0$. Hence, $\bar{A}\, u_2 \leq 0$ and $u_2\, e < 0$ which is contrary to our hypotheses. ⊠

Lemma 11: If X is a cell and $Au \leq 0$ with $eu < 0$ has no solution, then there are numbers η and ξ for which $eu \geq \eta \|v\| + \xi$ for all (u,v) in X.

Proof: Consider the set (u,v,w,z) such that (u,v) is in X, $w = eu$, $z \geq v_i$, and $z \geq -v_i$ for all i. Use Dines-Fourier-Motzkin elimination to reduce the system to inequalities in two variables w and z. Among these inequalities there must be at least one of form $w \geq \eta z + \xi$. ☒

The next result can be found in Karamardian [8].

Lemma 12: If X is convex and compact and $g : X \to R^n$ is continuous, then (X,g) has a stationary point.

Proof: $\pi(x - g(x))$ is a continuous map from X to X and has a fixed point. ☒

Lemma 13: If X is closed and convex, there are numbers η and ξ for which $eu \geq \eta \|v\| + \xi$ over X, and f is strongly monotone, then (X,e,f) has a stationary point.

Proof: Let X_t be the set of x in X such that $\|x\| \leq t$ for $t = 1, 2, 3, \ldots$. Select $(\bar{u},\bar{v})$ in some X_r. For $t \geq r$ let (u_t,v_t) be a stationary point of (X_t,e,f). If the v_t's have a cluster point v_* as $t \to +\infty$, then v_* is a v-stationary point of (X,e,f). Thus, suppose that $v_t \to \infty$ as $t \to +\infty$. We have

$$(u_t - \bar{u})\, e + (v_t - \bar{v})\, f(v_t) \leq 0$$

and hence

$$\alpha\|v_t - \bar{v}\|^2 \leq (v_t - \bar{v})(f(v_t) - f(\bar{v}))$$

$$\leq (\bar{u} - u_t)\, e + (\bar{v} - v_t)\, f(\bar{v})$$

and hence

$$\|v_t - \bar{v}\| \leq \frac{(\bar{u}-u_t)e}{\|v_t-\bar{v}\|}\,\alpha^{-1} + \|f(\bar{v})\|\,\alpha^{-1}$$

and hence for all sufficiently large t

$$\|v_t - \bar{v}\| \leq \frac{-u_t e}{\|v_t\|}\,\nu + \rho$$

for some ν and ρ that are invariant with t. By assumption the last expression is bounded and this contradicts our supposition that $v_t \to \infty$. ⊠

We don't know if the PIES model has a strongly monotone $\mathcal{Q}$ or $\mathcal{P}$, however, if it does the previous result asserts that the PIES model has a solution. Note that Lemma 13 implies Lemma 10.

4. Continuity of G

Given $V = \{v : \exists u,\ Au + Bv \leq C\}$ and $f : V \to R^m$ which is C^1 with p.d. derivatives, we define $G : V \to V$ by setting $G(v)$ to be the v-stationary point of (X,e,fv). Our algorithm generates the sequence $G^k(v_0)$. In this section we show G to be continuous.

The next lemma is a standard result. Let g and h be functions from Y to Y where Y is in R^n. By $\|g - h\|$ we mean the supremum of $\|g(y) - h(y)\|$ for all y in Y.

Lemma 14: If g is continuous, Y is compact, and g has fixed points Y_q, then for every $\varepsilon > 0$ there is a $\delta > 0$ such that $\|g - h\| \leq \delta$ implies that all fixed points of h lie within ε of Y_q.

Proof: Let 2δ equal the min of $\|g(y) - y\|$ over the set $\{y \in Y : \|y - Y_q\| \geq \varepsilon\}$. For $\|y - Y_q\| \geq \varepsilon$ we have $\|g(y) - y\| \geq 2\delta$. So, $\|h(y) - y\| \geq \|g(y) - y\| - \|g(y) - h(y)\| \geq 2\delta - \delta = \delta > 0$. ⊠

We define the distance $d(S_1,S_2)$ between two sets S_1 and S_2 in R^n to be (the Hausdorff metric):

$$\max_{x\in S_1} \min_{y\in S_2} \|x - y\| + \max_{y\in S_2} \min_{x\in S_1} \|x - y\|$$

Lemma 15: Let $\bar{X}$ be a cell in $R^{\ell+m}$ and $f : V \to R^m$ be C^1 with p.d. derivatives. Let $\bar{G}(v)$ be the v-stationary point of $(\bar{X},e,fv)$, if it

exists. If $\bar{G}$ is well defined for any v, then it is well defined on V and is continuous.

<u>Proof</u>: If (u_1,v_1) is a stationary point of $(\bar{X},e,fv_0)$, then $eu < 0$ with $\bar{A}u \leq 0$ has no solution, so $\bar{G}$ is well-defined by Lemma 10. Select $(\hat{A},\hat{B},\hat{c})$ so that $\hat{X} = \{(u,v) \in \bar{X} : \hat{A}u + \hat{B}v \leq \hat{c}\}$ is bounded and $\hat{A}u + \hat{B}v < \hat{c}$ for all (u,v) with $\|(u,v) - (u_1,v_1)\| \leq 3\varepsilon$. Select $\varepsilon_1 \leq \varepsilon$ so that $\|v - v_1\| \leq \varepsilon_1$ implies $d(\hat{U}(v), \hat{U}(v_1)) \leq \varepsilon$ where $\hat{U}(v)$ is the set u that minimizes eu subject to (u,v) in $\hat{X}$. Using Lemma 14 for $\bar{X}$ and the function $g(u,v) = \hat{\pi}((u,v) - (e,fv_0(v)))$ where $\hat{\pi}$ is the projection to $\hat{X}$, select δ so that if $\|v_0 - v_2\| \leq \delta$, then each fixed point, say $(\hat{u}_3,v_3)$, of $\hat{\pi}((u,v) - (e,fv_2(v)))$ is within ε_1 of some fixed point of $\hat{\pi}((u,v) - (e,fv_0(v)))$, say $(\hat{u}_1,v_1)$. Since $\|v_3 - v_1\| \leq \varepsilon_1$ there is a u_3 in $\hat{U}(v_3)$ so that $\|u_3 - u_1\| \leq \varepsilon$; (u_3,v_3) is a stationary point of $(\hat{X},e,fv_2)$. But $\|(u_3,v_3) - (u_1,v_1)\| \leq 2\varepsilon$ so (u_3,v_3) is a stationary point of $(\bar{X},e,fv_2)$ and $\|v_3 - v_1\| \leq \varepsilon$. ⊠

Thus, if G is defined anywhere, it is defined everywhere and is continuous.

5. The Functions G_i and Mixing

Let φ_i be any closed face of X of any dimension and let X_i be the affine hull of φ_i. For v in V define $G_i(v)$ to be the v-stationary point of (X_i,e,fv), if it exists; recall that if (X_i,e,fv) has a v-stationary point, then it is unique. We proceed to show that the action of G_i is equivalent to that of Newton's method as in Section 2.

Define $H_i : R^\ell \times V \to R^n$ by $H_i(u,v) = \pi_i((u,v) - (e,f(v))) - (u,v)$ where π_i is the projection of X_i. π_i is an affine function and can be expressed in the form $\pi_i(x) = Px + p$. Hence, H is linear in u. Given (u_0,v_0) in X the Newton iterate is any solution (u_1,v_1) in R^n to

$$H(u_0,v_0) + H'(u_0,v_0)((u_1,v_1) - (u_0,v_0)) = 0 \qquad \text{or}$$

$$P((u_0,v_0) - (e,f(v_0))) + p - (u_0,v_0)$$

$$+ (P - P\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & f'(v_0) \end{array}\right) - I)((u_1,v_1) - (u_0,v_0)) = 0 \qquad \text{or}$$

$$P((u_1,v_1) - (e,fv_0(v_1))) + p - (u_1,v_1) = 0 \qquad \text{or}$$

$$\pi_i((u_1,v_1) - (e,fv_0(v_1))) - (u_1,v_1) = 0 .$$

That is to say, given v_0 in V, the Newton iterate is any (u_1,v_1)

which is a stationary point of (X_i, e, fv_0) . That is to say, the v_i portion of the Newton iterate is $G_i(v_0)$.

Now suppose that (u_*, v_*) in $R^\ell \times V$ is a stationary point of (X_i, e, f) . The v-stationary point of (X_i, e, fv_*), namely v_*, is unique. We can now apply the results of Section 2; we consider only the v_0 for which $G_i(v_0)$ is in V .

First, there is a $\delta > 0$ and $\beta < 1$ such that if $\|v_0 - v_*\| \le \delta$ then $\|G_i(v_0) - v_*\| \le \beta\|v_0 - v_*\|$. Second, β can be made arbitrarily small by making δ sufficiently small. Third, if f' is Lipschitz continuous, then there is a γ and $\delta > 0$ so that $\|G_i(v_0) - v_*\| \le \gamma\|v_0 - v_*\|^2$ for $\|v_0 - v_*\| \le \delta$.

Given v_0 let (u_1, v_1) be a stationary point of (X, e, fv_0) . *Then* (u_1, v_1) *is interior to some face* φ_i *of* X . We then see that

$$G(v_0) = G_i(v_0) = v_1$$

which is to say that G is a mixing of the G_i's or that G is a mixing of a finite number of Newton methods.

6. Local Convergence Rate of G

Let v_* be the v-stationary point of (X,e,f). We show that if v_0 is sufficiently near v_*, then $G^k(v_0)$ tends to v_0 at a superlinear rate, and if, in addition, f' is Lipschitz continuous, the rate is quadratic. Regard the vectors d and d_* in R^n as constant functions from R^n to R^n.

Lemma 16: Assume x_* is a stationary point of (X,d_*). There is an $\varepsilon > 0$ such that if $\|d - d_*\| \leq \varepsilon$ and x is a stationary point of (X,d), then x is in the face $\{y \in X : y\, d_* = x_*\, d_*\}$ of X.

Proof: Suppose the contrary and we have $d_k \to d_*$ and x_k in X so that $d_*\, x_* < d_*\, x_k$ and $d_k\, x_k \leq d_k\, x$ for all x in X. Select an infinite subsequence so that all x_k lie in the relative interior of one face of X. Letting $\bar{x}$ be any element in the relative interior of this face we have $d_*\, x_* < d_*\, \bar{x}$ and $d_k\, \bar{x} \leq d_k\, x$ for all x in X. Setting $x = x_*$ and taking the limit we get a contradiction. ⊠

Select any u_* in $U(v_*)$ and let φ_1 be the face $\{(u,v) \in X : (e,f(v_*))((u_*,v_*) - (u,v)) = 0\}$ of X and let $\varphi_1, \varphi_2, \ldots, \varphi_h$ be all faces of φ_1 that meet $(U(v_*),v_*)$. Of course, φ_1 contains $U(v_*) \times \{v_*\}$. Defining X_i and G_i as in Section 5, we observe that any point in $\varphi_i \cap (U(v_*), v_*)$ is a stationary point of (X_i,e,f) for $i = 1, \ldots, h$. According to

Section 5, select $\delta = (\min \delta_i) > 0$ and $\beta = (\max \beta_i) < 1$ so that for $\|v_0 - v_*\| \leq \delta$ and $G_i(v_0)$ in V we have $\|G_i(v_0) - v_*\| \leq \beta\|v_0 - v_*\|$ for $i = 1, \ldots, h$. *The factor* β *can be made* arbitrarily small by making δ sufficiently small. If f' is Lipschitz continuous then there is a γ so that

$$\|G_i(v_0) - v_*\| \leq \gamma\|v_0 - v_*\|^2$$

for sufficiently small δ .

Select $\Delta_1 \leq \delta$ so that if (u,v) is in φ_1 and within a distance Δ_1 of the set $U(v_*) \times \{v_*\}$ then (u,v) is in the relative interior of one of the faces $\varphi_1, \ldots, \varphi_h$. Select $\Delta_2 \leq (1/2)\Delta_1$ so that for $\|v - v_*\| \leq \Delta_2$ we have $d(U(v), U(v_*)) \leq (1/2)\Delta_1$. Select $\varepsilon > 0$ according to Lemma 16 for $d_* = (e, f(v_*))$. Select $\Delta_3 \leq \Delta_2$ so that for v_0 and v within a distance Δ_3 of v_* we have $\|fv_0(v) - f(v_*)\| \leq \varepsilon$. Select $\Delta_4 \leq \Delta_3$ so that for $\|v_0 - v_*\| \leq \Delta_4$ we have $\|G(v_0) - G(v_*)\| \leq \Delta_3$.

Now, for $\|v_0 - v_*\| \leq \Delta_4$ we have $\|G(v_0) - v_*\| \leq \Delta_3$. Consequently, $\|fv_0(G(v_0)) - f(v_*)\| \leq \varepsilon$. Selecting any u_1 in $U(v_0)$ we have $(u_1, G(v_0))$ in φ_1 . Since $\|G(v_0) - v_*\| \leq \Delta_2$ there is a point u_* in $U(v_*)$ with $\|(u_1, G(v_0)) - (u_*, v_*)\| \leq \Delta_1$. Consequently, $(u_1, G(v_0))$ lies interior to one of the faces $\varphi_1, \ldots, \varphi_h$. Therefore, $G(v_0) = G_i(v_0)$ for some $i = 1, \ldots, h$. Thus, $\|G(v_0) - v_*\| \leq \beta\|v_0 - v_*\|$. *Reusing the last expression we obtain* $G^k(v_0) \to v_*$. Since

β can be made arbitrarily small by making δ sufficiently small we obtain $\|G^{k+1}(v_0) - v_*\| \leq \beta_k \|G^k(v_0) - V_*\|$ where $\beta_k \to 0$ as $k \to \infty$. Similarly, if f' is Lipschitz continuous, we have

$$\|G^{k+1}(v_0) - v_*\| \leq \gamma \|G^{k+1}(v_0) - v_*\|^2$$

for $k \geq K$ for some K.

Note that if $\varphi_1 = U(v_*) \times \{v_*\}$, then $G^k(v_0) = v_*$ for some k, that is, the algorithm terminates.

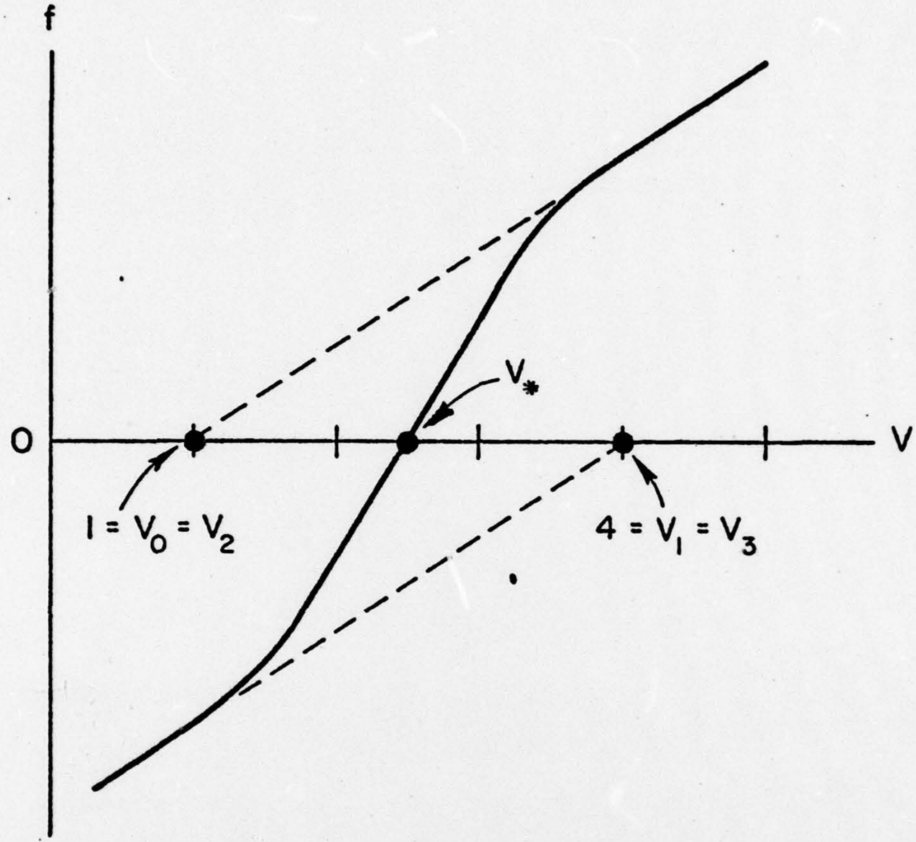
Finally, we observe that there is a η such that

$$d(U(v), U(v_*)) \leq \eta \|v - v_*\| .$$

Thus we have

$$d(U(G^k(v_0)), U(v_*)) \leq \eta \|G^k(v_0) - v_*\|$$

As the figure illustrates, the algorithm of the paper is not globally convergent; let $\ell = 0$, $m = 1$, $X = [0,5]$.



f
0
V
V_*
$1 = V_0 = V_2$
$4 = V_1 = V_3$

7. Bibliography

[1] AHN, B-h, Computation of Market Equilibrium for Policy Analysis: The Project Independence Evaluation System Approach, Department of Engineering-Economic Systems, Stanford University, (February 1978)

[2] COTTLE, R. W., The Principal Pivoting Method of Quadratic Programming, Mathematics of the Decision Sciences, Part 1, Am. Math. Soc., Eds: G. B. Dantzig and A. F. Veinott, Jr.,(1968), 144-162.

[3] EAVES, B. C., Computing Stationary Points, Math. Prog. Study 7, (1978), 1-14.

[4] HAN, S-P, Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems, Math. Prog. 11 (1976), 263-282.

[5] HOGAN, W. W., Energy Policy Models for Project Independence, Comput. & Ops. Res. 2 (1975), 251-271.

[6] ____________, Project Independence Evaluation System: Structure and Algorithms, Proceed. of Symposia in Appl. Math. 21 (1977).

[7] IRWIN, C. L., Convergence Properties of a PIES-Type Algorithm for Non-Integrable Functions, TR SOL 77-33, Department of Operations Research, Stanford University, (December 1977).

[8] KARAMARDIAN, S., The Complementarity Problem, Math. Prog. 2 (1972), 107-129.

[9] LEMKE, C. E., Bimatrix Equilibrium Points and Mathematical Programming, Management Sciences 11 7 (1965), 681-689.

[10] MATHIESEN, L., Efficiency Pricing in a Linear Programming Model: A Case with Constraints on Dual Variables, TR SOL 74-18, Department of Operations Research, Stanford University, (November 1974).

[11] ____________, A Linear Complementarity Approach to General Equilibrium, WP SOL 74-9, Department of Operations Research, Stanford University, (December 1974)

[12] ____________, Two Equivalent Programming Models for Marginal Cost Pricing, WP SOL 75-1, Department of Operations Research, Stanford University, (February 1975).

[13] ORTEGA, J. M., Numerical Analysis (1972), Academic Press.

[14] PALOMARES, U.M.G., and O. L. MANGASARIAN, Superlinearly Convergent Quasi-Newton Algorithms for Nonlinearly Constrained Optimization Problems, Math. Prog. 11 (1976), 1-13.

[15] POWELL, M.J.D., A Fast Algorithm for Nonlinearly Constrained Optimization Calculations, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, (June 25, 1977).

[16] ______________, The Convergence of Variable Metric Methods for Nonlinearly Constrained Optimization Calculations, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, (August 5, 1977).

[17] ROBINSON, S. M., A Quadratically-Convergent Algorithm for General Nonlinear Programming Problems, Math. Prog. 3 (1972), 145-156.

[18] ______________, Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear-Programming Algorithms, Math. Prog. 7 (1974), 1-16.

[19] ROCKAFELLAR, R. T., Monotone Operators and Augmented Lagrangian Methods in Nonlinear Programming, Department of Mathematics, University of Washington, Seattle. To appear in Nonlinear Programming 3, Eds: O. L. Mangasarian, et al, Academic Press (1978).

[20] TAPIA, R. A., Diagonalized Multiplier Methods and Quasi-Newton Methods for Constrained Optimization, Optimization Theory and Applications 22 2 (June 1977), 135-194.

[21] ____________, Quasi-Newton Methods for Constrained Optimization: Equivalence of Existing Methods and a New Implementation, Department of Mathematical Sciences, Rice University, (July 1977).

[22] THRASHER, R. L., Notes on Convergence of PIES-like Iterative Processes, New Mexico State University, (February 23, 1978).

[23] TODD, M. J., New Fixed-Point Algorithms for Economic Equilibria and Constrained Optimization, TR 362, School of Operations Research and Industrial Engineering, Cornell University, (September 1977).

[24] TSE, E., Monotonic Descent Algorithms for Finding the Market Equilibrium, Department of Engineering-Economic Systems, Stanford University.

[25] WILSON, R. B., A Simplicial Algorithm for Concave Programming, D.B.A. Thesis, Harvard University (May 1963, Revised May 1964).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ~~SOL 78-13~~	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) A LOCALLY QUADRATICALLY CONVERGENT ALGORITHM FOR COMPUTING STATIONARY POINTS,		5. TYPE OF REPORT & PERIOD COVERED (9) Technical Report,
		6. PERFORMING ORG. REPORT NUMBER (14) SOL-78-13
7. AUTHOR(s) (10) B. Curtis Eaves		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-78-G-0026, EY-76-S-03-0326
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research -- SOL Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office Box CM, Duke Station Durham, N.C. 27706 (12) 31 p.		12. REPORT DATE (11) May 1978
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(18) ARO (19) 15254.2-M

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Local Convergence	Variational Problems
Quadratic Convergence	Newton's Method
Stationary Points	Complementary Theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Stationary points for a system with a convex polyhedral set and a continuously differentiable function with positive definite derivatives are computed by iteratively solving the linearized problem. The procedure is shown to be a mixing of a finite number of Newton methods and to have a local convergence rate which is quadratic. The stationary point problem is of the type arising from the PIES energy model.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408 765